



Al-Samman, I., Doufexi, A., & Beach, M. (2017). A Proposal For Hybrid SDN C-RAN Architectures for Enhancing Control Signaling Under Mobility. In *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall): Proceedings of a meeting held 18-21 September 2016, Montreal, Quebec, Canada* Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/VTCTFall.2016.7880872>

Peer reviewed version

Link to published version (if available):
[10.1109/VTCTFall.2016.7880872](https://doi.org/10.1109/VTCTFall.2016.7880872)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via IEEE at <http://ieeexplore.ieee.org/document/7880872/>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

A Proposal For Hybrid SDN C-RAN Architectures for Enhancing Control Signaling Under Mobility

Imad Al-Samman, Angela Doufexi and Mark Beach

Communication Systems & Networks Group, Department of Electrical and Electronic Engineering
University of Bristol, Bristol, United Kingdom

Email: {[Imad.Al-samman](mailto:Imad.Al-samman@bristol.ac.uk), [M.A.Beach](mailto:M.A.Beach@bristol.ac.uk), [A.Doufexi](mailto:A.Doufexi@bristol.ac.uk)}@bristol.ac.uk

Abstract- The vast growth in mobile data traffic requires future mobile network infrastructure to support fast overall growing mobile devices while improving the cost and the energy efficiencies. Many recent studies in this context have focused on proposing solutions for end-to-end architectural designs based on flexible allocations of functions, network function virtualization and software-defined implementations (SDN). It's also of upmost importance to introduce a 5G End-to-End architecture that support low latency RAN with joint design of programmable and software's driven networks that can adapt dynamically to the fluctuating traffic demands. This study contributes in this area by providing innovative joint Cloud-RAN-SDN network designs which are expected to reduce operational cost and overall control signaling load. The paper proposes two architectural schemes called C-RAN.C-MME and C-RAN.D-MME. We demonstrate the associated signaling load analysis, evaluate the performance and compare it against existing literature architectures. Evaluation results indicate a substantial improvement in terms of reducing the signaling load taking into account several network metrics. It's worth mentioning that this study shows the impact of the proposed architectures on the signaling plane only, while their benefits for the transmission/data planes is considered in future work.

Key Words—C-RAN; SDN; LTE, Control signalling.

I. INTRODUCTION

The exponential increase in mobile networks connected devices will lead to a data tsunami in the coming years. According to [1], it is expected that the data transmission volume will grow 10 folds by 2019. This will require the mobile networks with its current shape to cope with an unprecedented rate of growth in network usage. Moreover a study in this context refers to the end of an era for mobile network profitability in 2015 [2]. The reason behind the revenue decrease is the cost of expanding and upgrading the traditional decentralised architecture to meet the mobile traffic surge while the revenue is not growing at the same rate. The increase of smart phone applications along with their related keep-alive signalling would cause a major challenge for operators in respect of increasing the load of LTE signalling to keep up with every short message generated by them. In that context, the operators need to come up with new approaches to face the aforementioned difficulties. The emergence of software define networking (SDN) [3] in wired networks through its OpenFlow (OF) protocol which enables the SDN controller to directly interact with forward plane of network devices such as switches and routers. SDN attractive features in separating the data and control planes have inspired the researchers in both academia

and industry. The SDN approach is based on the separation between the control and the data planes basically in the Evolved Packet Core (EPC) which consists of a Mobility Management Entity (MME), Serving Gateway (SGW) and PDN Gateway (PGW), while the E-UTRAN is still responsible of managing the radio functions with the end user. In this context there have been many recent studies which have brought the advantages of SDN utilization in LTE into the light [4][5][6]. However the previous studies have only considered the D-RAN (Distributed Radio Access Network) architecture. This study will investigate the potential performance gain of utilising SDN in LTE network with a Cloud RAN (C-RAN) topology and compare it against the D-RAN related studies. The authors in [7] have examined the importance of using C-RAN in a handover context, where the signalling overhead is expected to increase due to the deployment of multi-tier cellular networks. Therefore taking C-RAN into account along with the afore-cited SDN approaches will help in terms of reducing the overall signalling and simplifying the network topology from the controller perspective.

In this paper we propose new architectural schemes based on what it is stated above and perform their overall signalling load analysis. The objective is to determine a better architecture in terms of lower signalling load. The related analysis will address multiple network parameters such as user density velocity and tracking area update. The rest of this paper is organised as follows. Section II describes the calculation of maximum distance between RRH (Radio Remote Head) and BBU (Base Band Unit). The new proposed network architectures with the corresponding mathematical modelling of different network metrics is introduced in section III. Section IV evaluates the performance of the new schemes and compares them against literature schemes. Finally we conclude the paper in section V.

II. C-RAN SIZE "PROBLEM FORMULATION"

Despite the benefits of using C-RAN, the main challenge is the maximum radius of a C-RAN that a BBU can manage. The UE in a standard LTE network should receive ACK/NACK from eNB in three sub-frames after sending uplink data to comply with the HARQ protocol. Hence the eNB needs to finish the DL processing within 3 ms after receiving the UL data. In C-RAN this timing requirement is difficult to meet due to the transmission line of fibre optics. This new medium links the RRH and the BBU which functions as the eNB. In addition, there is a processing delay introduced by the active equipment

in the front-haul links. According to [8], in order to meet the timing constraint in the standards the vendors need to amend the BBU design to accelerate the DL process and shorten it to 2.7 ms. The aim of the delay reduction is to compensate the latency that occurs due to the newly added separation distance between the RRH and the BBU. Overall transmission delay in the UL and its ACK/NACK response is presented in [7]. Table I illustrates the required delay values at all network components mentioned previously to satisfy the overall delay requirements of 3 ms. The Max fibre Round Trip Time (RTT) is given by: $3 \text{ ms} - (A + B + C + D) = 246 \mu\text{s}$ where A, B, C and D are defined in Table I. By taking into account that the transmission latency on fibre is 5 $\mu\text{sec/km}$, the Max fibre distance = $\frac{\text{Max Fibre RTT}}{2 \times 5 \mu\text{sec/km}} = 24.6 \text{ km}$. we assume the standard cell radius r_c in LTE is 4 km through the rest of the study and a hexagonal shape of cells for both RRH and the C-RAN itself in D-RAN and C-RAN. The area of hexagonal based RRH then can be calculated as $A = \frac{r_c^2 \times 3 \times \sqrt{3}}{2} = 24\sqrt{3} \text{ km}^2$ for C-RAN which its radius equals to 24 km according to previous analysis, hence $A_c = 36 A$. This leads to each C-RAN to be composed of 36 RRHs. On the other side, the area of hexagonal based C-RAN is $A_c = \frac{R_c^2 \times 3 \times \sqrt{3}}{2}$.

TABLE I
DELAY COMPONENT VALUES

Delay component	Unit of delay	Required values
A. RT RF processing time	RRH	~ 40 μsec
B. RT CPRI processing time	RRH,BBU	~ 10 μsec
C. RT BBU processing time	BBU	~ 2700 μsec
D. Front-haul equipment's (if exist) delay	Front-haul	~ 4 μsec

III. PROPOSED C-RAN ARCHITECTURES

This section highlights the relevant studies in the area and their findings. That's besides clarifying the proposed architectures in subsections B and C with their control signalling analysis.

A. Related Work

The authors in [4] [5] have adopted a partial SDN approach based on decoupling the control and data planes at the serving gateway (SGW) which becomes an advanced OF switch that enables to encapsulate/ de-capsulate GTP packets while SGW-C has been transferred to the entity where the OF controller and MME reside. This approach is innovative and has shown an improvement in the total reduction of signalling load. However the new design has been considered as a partial solution as the PDN gateway (PGW) as the functional entity is following the existing 3GPP architecture. The authors in [6] have proposed a complementary vision which is fully realized in Openflow where the PGW-C has been decoupled and virtualized as an application running on top of the OF controller. This methodology is based on a complete separation between the

control and data planes. For convenience, the proposed architectures in [6] are called OF D-RAN and partial OF D-RAN in [4][5]. Five main signalling procedures are evaluated in previous studies, taking into consideration the handover and tracking Area Update. The analytical results of adopting OF D-RAN against the legacy topology & partial OF D-RAN show an overall improvement in decreasing the number of messages exchanged between all entities in an hour. It has been assumed that each UE supports multiple applications (K types) with certain arrival rate of λ_k , in addition, hexagonal shaped cells of number C have been chosen and RRH area of consideration A, N_{ue} is total number of UEs and ρ_{ue} is the user density given by $\frac{N_{ue}}{A}$.

B. Proposed Archetype 1

The proposed architectures will analyse the signalling load based on the signalling analysis model projected in [9]. The first architecture applies major amendment on legacy and literature architectures and utilizes the methodology of Distributed MME signalling load. For later ease of use we will call it C-RAN.D-MME. This architecture is based on integration of the MME functionality within the BBU in the data, in this aspect, each C-RAN is seen as single cell with its own BBU and MME pooling. One C-RAN is composed of variable number of RRH. The SGW-C, PGW-C and OF controller are combined and packaged in one entity called Centric Controller CC. Three main C-RANs are proposed as illustrated in Fig.1. [4], [5] and [6] evaluations have addressed MME signalling only as they count only the incoming and departing MME messages, while this study will evaluate all messages flow between all control entities. Fig. 2 demonstrates the call flow for C-RAN.D-MME initial attachment. It is proposed to register UE information and apply authorisation and authentication procedure plus creating sessions between the MME and SGW-C. The call flow is aligned with [10]. The total number of messages is equal to 8 (RRC connection reconfiguration and reconfiguration complete are considered as one message), where the probability of UE initiates an attachment procedure in the network is assumed $P_{initial} = 0.2$, and thus the total signalling load $S_{initial}$ is:

$$S_{initial} = P_{initial} \cdot A \cdot \rho_{ue} \cdot C \cdot 8. \quad (1)$$

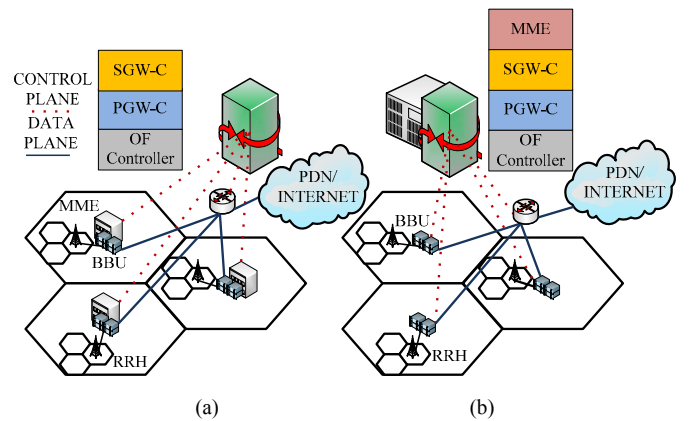


Fig. 1. Proposed CRAN with (a) Distributed MME & (b) Centralised MME.

The second procedure to address is the UE-generated services. This event takes place when the UE which is in IDLE

state triggers the need to set up a connection with the PDN/Internet. In the standard LTE architecture, a few messages are exchanged between the eNB and the MME including the authentication check, initial context setup request and initial context setup response. However C-RAN.D-MME eliminates the need for such messages. The BBU will be assessed as the eNB in former studies (OF enabled entity) thus whatever applies for the OF switch can be mapped in the same manner on it. That's why when the UE sends its first packet to the BBU, the BBU looks up in its flow table to find a matching rule. In case the entry is not found, an OF (packet-in) message is sent to the CC in order to investigate the message, acquire source and destination IP address and interact with SGW-C and PGW-C in an internal process to obtain the GW-U required information. Consecutively, the CC generates flow rules for subsequent packets. Each service k has an arrival rate λ_k (sessions/hour/UE) and average session duration of μ_k^{-1} . We denote P_k as the probability that session k is generated by the UE, The number of total messages in this case is 6, therefore, the total signalling for such an event is given by:

$$S_{uetrig} = \lambda_k \cdot P_k \cdot \rho_{ue} \cdot C \cdot A \cdot 6 \quad (2)$$

The third procedure is network triggered service, where in this analysis both cases of the UE being either in IDLE or CONNECTED state are taken into account. C-RAN.D-MME new architecture proposes some changes in the entities functionalities. Unlike the paging procedure in the standards [9], where paging the IDLE UEs is one of the MME functions, in our proposed architecture we are assuming that related connection management is one of CC's functions instead of MME. This implies when an incoming session triggers the CC, it will page all MMEs in its domain. This leads to unicast paging for a limited number of times (the same as number of C-RAN in its domain) instead of paging all eNBs. When the UE is in CONNECTED state, the number of exchanged messages is reduced by one, which is the paging. The total signalling load can be given by:

$$S_{NWtriggered} = ((8 + C_{ran}) \cdot R_p \cdot P_i + 7 \cdot (1 - P_i)) \cdot \lambda_k \cdot (1 - P_i) \cdot \rho_{ue} \cdot C \cdot A \quad (3)$$

Where R_p is the average number of paging transmission per page, C_{ran} is the number of C-RAN cells in considered area. P_i is the probability of UE being in IDLE state. It can be computed from the process of (X_n, Y_n) which presents CONNECTED and IDLE states of type-n session respectively that $E[X_n] = \mu_n^{-1}$ and $E[Y_n] = \lambda_n^{-1}$ according to the alternating renewal process and independence assumption of applications [9]. P_i can be computed as $P_i = \prod_{n=1}^N \frac{\mu_n}{(\lambda_n + \mu_n)}$, the UE is CONNECTED state when it has at least one active session thus its related probability is: $P_{co} = 1 - P_i$.

The fourth procedure is the handover (HO), for simplicity the HO will be divided into two parts: Firstly inner handover within the CRAN itself, it only considers the mobility within the boundary of CRAN. Secondly outer handover on other side occurs between CRANs themselves. Inter-technology handovers will be ignored in this paper as the concentration is on LTE only. In LTE, the handover process can be done via two interfaces, either by X_2 between eNBs or via S_1 which depends mainly on MME. Inner handover is calculated based on X_2 or

S_1 while outer handover is based only on S_1 . Inner handover in C-RAN.D-MME doesn't experience inter-MME or inter-GWs as in legacy topology and [9] analysis. Fig.3 illustrates the call flow when $X_2 - HO$ triggers the network. In this architecture we can save signalling caused by messages exchanged between source and target eNBs such as HO request, HO Ack, SN status transfer, and between target eNBs and MME such as Path switch request and Path Switch Request Ack.

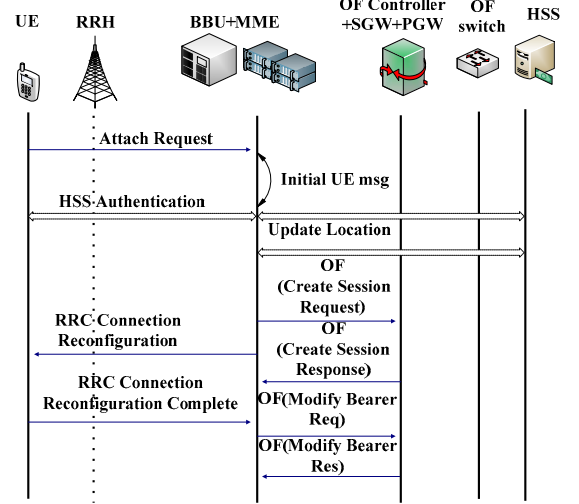


Fig. 2. Architecture 1 initial attachment call flow

In addition, the CC notifies the OF switch to modify its flow table by sending OF_Packet_Out message and communicates with D-MME to exchange the RRH address and TEID for downlink user plane via OF_Packet_in (modify Bearer Request) and OF_Packet_Out (Modify Bearer Response) as in [10]. Based on fluid flow model [11], the mobile crossing rate out of an enclosed area of perimeter length L can be given by $R_c = \frac{\rho_{ue} \cdot L \cdot V}{\pi}$. Where V is the UE average velocity. L is the cell perimeter length. From Fig.3 it can be observed that four main messages are exchanged (Radio Resource Control (RRC) Reconfiguration and RRC Reconfiguration complete are considered as one message), hence X_2 based inner HO signalling load is computed by:

$$S_{1hx2cran} = R_c \cdot (1 - P_i) \cdot C_{inran} \cdot C_{ran} \cdot 4 \quad (4)$$

C_{inran} is the RRHs number in single C-RAN, C_{ran} is number of C-RAN cells. The multiplication $C_{inran} \cdot C_{ran}$ results in total number of cells in the considered area. Outer HO calculation is assumed to be based only on S_1 interface between C-RANs, crossing rate out in outer HO is based on L_{cran} (the perimeter of single C-RAN). Hence the outer HO calculation is given by:

$$S_{1hos13} = R_{cran} (1 - P_i) C_{ran} 9 \quad (5)$$

Since each C-RAN has its own MME, The HO that occurs between C-RANs can be considered as inter-MME HO, however we have only one SGW-C in our architecture, consequently there is no SGW-C relocation. Details of signalling procedures are not presented in this work as the purpose is showing the advantages of reshaping the network architecture rather than discussing signalling messages which are presented in details in [10].

The final important procedure is Tracking Area Update (TAU). This event is initiated when the UE moves and detects a new tracking area that is not in the tracking areas list allocated by the MME at the time of UE attachment or when the TAU timer expires. However we will ignore the second condition in this study. This procedure has different call flows in legacy LTE/EPC architectures depending on MME relocation or not and arises irrespective of whether the UE is in IDLE or CONNECTED state. It has one constant call flow in this proposed scheme, C-RAN.D-MME assumes that each C-RAN is a tracking area by itself served by its MME, no TAU occurs unless the UE crosses between C-RANs. The UE sends TAU to the RRH which passes it to the BBU. The MME is integrated and located at the same data centre as the BBU thus no signalling is needed between them. The target MME will then update the location of the UE to the home subscriber server (HSS) for future incoming sessions. A couple of messages are required to be exchanged (source and target MME) about cancelling the location information at the source MME and inserting subscription data at the target one. When the UE is in CONNECTED state the target MME has to communicate with the SGW-C for U-plane programming. The rate of crossing the tracking area is estimated by crossing out of a cell multiplied by $\frac{1}{\sqrt{C_{inran}}}$ where $\sqrt{C_{inran}}$ is the tracking area size in this scheme. Concluding above the total signalling for TAU is:

$$S_{taudist} = \frac{1}{\sqrt{C_{inran}}} \cdot R_c \cdot C \cdot 10. \quad (6)$$

C. Proposed Archetype 2

The second proposed architecture (C-RAN.C-MME) resembles the first scheme in terms of C-RAN topology as demonstrated in Fig1(b), nevertheless it considers a central MME as a pool of MMEs rather than distributed, the functions of the central MME are virtualised as an application like SGW-C, PGW-C in C-RAN.D-MME where all of them run on top of OF controller and communicate with it through API. The rest of the network entities are kept the same. By initial attachment, the UE sends attach request message to the BBU which inserts it in the Openflow attach request and sends it to the CC. IP allocation will be performed by PGW-C application, while the MME application triggers the authentication process, then attach accept is embedded in Openflow initial context setup request message and sent from the CC to the BBU. The initial attachment total signalling in this scenario is given by:

$$S_{initial} = P_{initial} \cdot A \cdot \rho_{ue} \cdot C \cdot 8. \quad (7)$$

Fig.4 illustrates the UE-triggered service call flow in C-RAN.C-MME that has slight changes as compared to the C-RAN.D-MME. The UE starts the procedure by sending service request to the BBU, the message is then embedded in open flow message and sent again to the CC, MME application processes the authentication part, prepares the resources and initiates "Openflow initial context setup request" message to the BBU. After the radio configuration, the UE sends its first packet to the BBU which insert "OF-packet-in" in "OF initial context response" to the CC, the CC then interacts with SGW-C and PGW-C for IP allocation. UE-triggered total signalling is given then by:

$$S_{uetrig} = \lambda_k \cdot P \cdot \rho_{ue} \cdot C \cdot A \cdot 8. \quad (8)$$

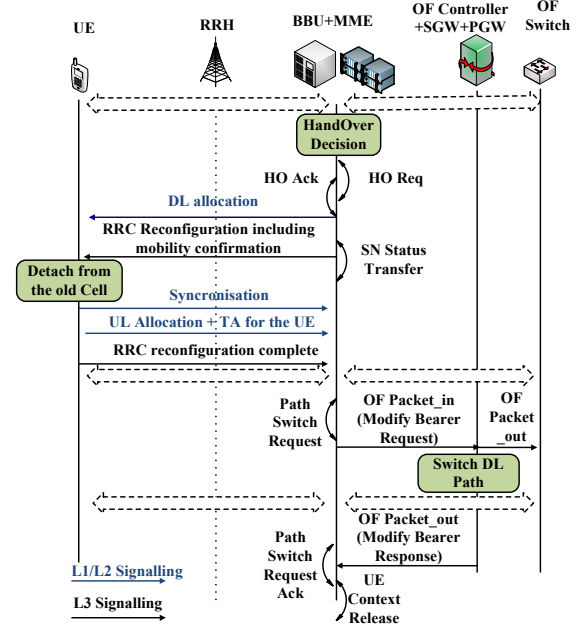


Fig. 3. Architecture 1 X2 HO call flow

The C-RAN.C-MME has only one central MME and the size of the tracking area is not constant but variable due to its configuration. When DL traffic needs to be delivered to the UE in the IDLE state (only its tracking area known to MME) the CC performs paging and sends unicast messages to all eNBs (RRHs) in its tracking area. The total signalling load for this procedure is given by:

$$S_{NWtriggered} = [((10 + C_{tau}) \cdot R_p \cdot P_i + 9 \cdot (1 - P_i)) \cdot \lambda_k \cdot (1 - P) \cdot \rho_{ue} \cdot C \cdot A]. \quad (9)$$

Where C_{tau} is the number of RRHs per tracking area. Handover in C-RAN.C-MME is divided the same manner into inner and outer parts as C-RAN.D-MME, Total signalling for X2 based inner HO is given by:

$$S_{1hx2cran} = R_c \cdot (1 - P_i) \cdot C_{inran} \cdot C_{ran} \cdot 4. \quad (10)$$

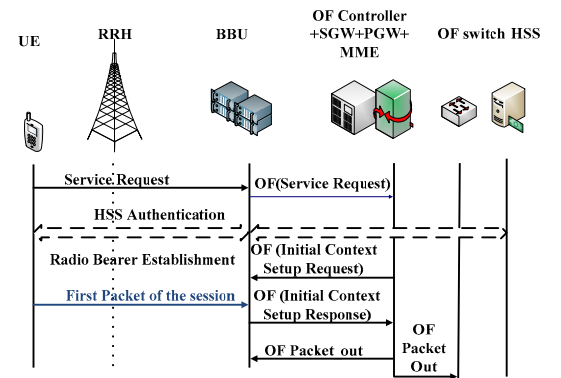


Fig.4. UE-triggered service call flow

While total signalling due to outer HO:

$$S_{lhos13} = R_{cran} \cdot (1 - P_i) \cdot C_{ran} \cdot 9. \quad (11)$$

As there is only one virtualised MME in the CC, the tracking area update procedure is intra-MME only. In this architecture, the MME simply records the UE new location and accepts the TAU, therefore there is no signalling to the HSS. Total signalling for TAU procedure is given:

$$S_{ltaudist} = \frac{1}{\sqrt{C_{tau}}} \cdot R_c \cdot C \cdot 6. \quad (12)$$

Tracking area update signalling is different to (C-RAN.D-MME) as it depends on variable parameter C_{tau} .

IV. NUMERICAL AND ANALYTICAL RESULTS

In this section, we present the numerical results of the signalling load for both proposed (C-RAN.C-MME) & (C-RAN.D-MME) architectures, legacy network and OF D-RAN architecture proposed in [6]. The calculations in this paper will only acknowledge one paging case which is unicast (uni) as referred in 4, 5 & 6. [6] and [9] have investigated the optimality between multicast and unicast, thus it is sufficient in this study to consider just one of them. The evaluation will consider various metrics related to either the network or users. In this study, it is presumed that each user is experiencing 3 different applications with their related service time as stated in Table II.

TABLE II USERS APPLICATIONS PARAMETERS		
Application	Session Arrival Rate	Session time
Voice	0.3	0.1
Chat	0.4	0.01
Web browsing	0.05	0.05

Other assumptions are made such as $P_k = 0.5$, $R_p = 1.1$. Uniform hexagonal cells are deployed with an overlapping factor of $\gamma = 1.2$. The first analytical case is based on varying the user's velocity. It is intuitive that increasing the velocity will increase the overall signalling load as the two main procedures (HO & TAU) will occur more frequently regardless of the deployed architecture. The total signalling load is calculated by basic summing of (1, 2, 3, 4, 5, 6) equations and (7, 8, 9, 10, 11, 12) for C-RAN.D-MME, and C-RAN.C-MME, respectively. The study assumes that three quarters of the users are experiencing mobility with variable speeds as shown in Fig.5. The figure illustrates that C-RAN.D-MME (scenario1) results in the least signalling load, followed by C-RAN.C-MME (scenario 2) and the OF D-RAN. However C-RAN.D-MME is not always the most efficient architecture to implement as other network factors can significantly impact the presented results such as the TA (Tracking Area) size which is referred to as C_{tau} . The conclusions drawn from Fig.5 are applicable for the case of small TA size ($C_{tau} = 6$), where users in C-RAN.C-MME and OF D-RAN need to perform TAU procedure more frequently than C-RAN.D-MME since the latter is only dependent on the number of RRHs per CRAN cell. Nevertheless when the TA becomes greater ($C_{tau}=36$), results show leading deviation as in Fig.6. Applying the C-RAN.D-MME architecture will save up to 68% of signalling load when all users are in a stationary state, however when users experience mobility, the saving decreases

to 31% compared to our legacy network. On the other hand, C-RAN.C-MME only results in 11% saving in signalling load compared to the legacy architecture when users are in a stationary state. As we increase the user's velocity, the C-RAN.C-MME results in greater signalling load reduction of about 40% when the velocity reaches 100km/hr. Furthermore, the C-RAN.C-MME attains best performance compared to all other architectures when the speed is beyond 65km/hr. It is observed that the performance of the literature architecture OF D-RAN is bounded by our proposed and legacy architectures. It is worth mentioning that at very high speed the performance of the C-RAN.D-MME converges to the OF D-RAN. Another metric to consider is the number of users in the tested area, analysis can be projected over a range of [500000 ~ 3000000] users while keeping the area invariable. Fig.7 highlights the differences between the proposed and literature architectures load. It is observed that C-RAN.D-MME results in the best performance in terms of overall signalling reduction across all simulated user's densities when C_{tau} is only 6 RRHs.

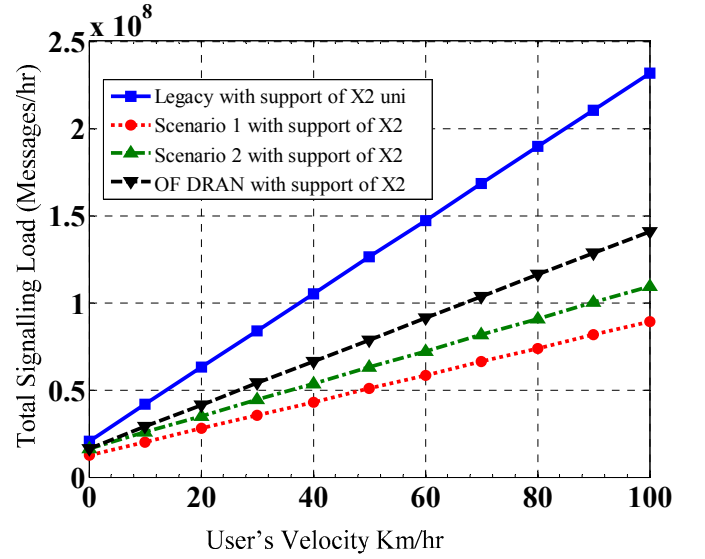


Fig. 5. Velocity based Comparison between Scenario 1 & 2 $C_{tau}=6$

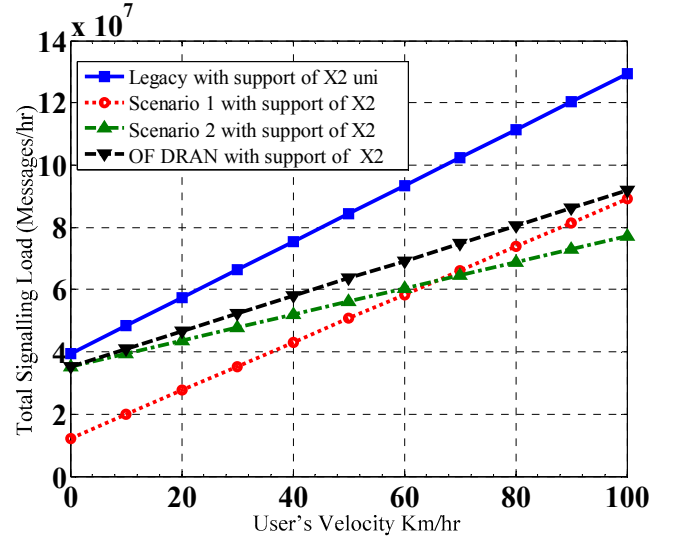


Fig. 6. Velocity based Comparison between Scenario 1 & 2 $C_{tau}=36$

In the same manner, these observations are not applicable for the case of higher TA size ($C_{tau}=36$). For the same number of RRHs within the TA size for both proposed architectures, the users in both C-RAN.C-MME and C-RAN.D-MME demonstrate the same rate of TAU. When user's density is in the range of 500000 users in the considered area, the C-RAN.D-MME, C-RAN.C-MME ($C_{tau}=36$) and OF D-RAN ($C_{tau}=36$) experience almost the same signalling load. However when the user density increases, C-RAN.C-MME ($C_{tau}=36$) shows better performance in terms of the lower signalling load. OF D-RAN ($C_{tau}=36$) and C-RAN.D-MME experience almost the same signalling load but not at high density levels. Table III illustrates the saving ratios for each architecture with different C_{tau} values which is labelled as T_{au} in Fig.7. The saving ratios results are independent of user's density change.

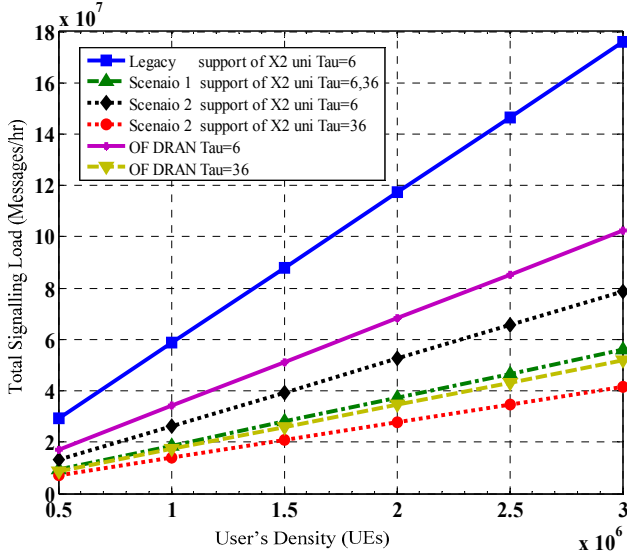


Fig. 7. UE's Density based comparison between Sc1 & Sc2, $C_{tau}=6, 36$

TABLE III
SAVING RATIOS FOR PROPOSED AND LITERATURE ARCHITECTURES

Architecture Type	Saving ratio $C_{tau}=6$	Saving ratio $C_{tau}=36$
C-RAN.D-MME (Sc1)	68%	68%
C-RAN.C-MME (Sc2)	55%	76%
OF D-RAN	41%	70%

V. CONCLUSION

In this paper we proposed new C-RAN architectures with central and distributed MMEs that combine a SDN LTE architecture with a C-RAN. Two design schemes have been proposed which are: C-RAN.C-MME and C-RAN.D-MME. The paper has considered different users parameters such as velocity and user's density. In addition, the paper analysed the impact of the TA size (C_{tau}) metric as a network parameter on system performance. The analytical results show that for different range of velocities and user's densities, C-RAN.D-MME performs better for small C_{tau} , while C-RAN.C-MME

outperforms all architectures including the legacy for cases of larger TA size ($C_{tau}=36$) in terms of the lowest signalling overhead. The paper presented the impact of the proposed C-RAN architectures on the signalling load and didn't investigate their impact of the data plane, which is part of our future work.

REFERENCES

- [1] "Cisco visual networking index: Global mobile data traffic forecast update, 2015-20" 3 Feb 2016.
- [2] Tellabs, "Tellabs "end of profit" study executive summary," Tech. Rep., January 2011.
- [3] Open Networking Foundation (ONF), Software-defined networking: The new norm for networks. ONF White paper (2012). [Online] <https://www.opennetworking.org/images/stories/downloads/sdn-resources/white-papers/wpsdn-newnorm.pdf>.
- [4] S. Ben Hadj Said *et al.*, "New control plane in 3GPP LTE/EPC architecture for on-demand connectivity service," *Cloud Networking (CloudNet), 2013 IEEE 2nd International Conference on*, San Francisco, CA, 2013, pp. 205-209.
- [5] M. R. Sama, S. Ben Hadj Said, K. Guillouard and L. Suci, "Enabling network programmability in LTE/EPC architecture using OpenFlow," *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2014 12th International Symposium on*, Hammamet, 2014, pp. 389-396.
- [6] V. G. Nguyen and Y. Kim, "Proposal and evaluation of SDN-based mobile packet core networks," *J Wireless Com Network EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, 2015.
- [7] L. Liu, F. Yang, R. Wang, Z. Shi, A. Stidwell and D. Gu, "Analysis of handover performance improvement in cloud-RAN architecture," *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*, Kun Ming, 2012, pp. 850-855.
- [8] Harrison J. Son and S.M.Shin. (April 2014). Fronthaul Size: Calculation of maximum distance between RRH and BBU.[online] <http://www.netmanias.com/en/?m=view&id=blog&no=6276>.
- [9] I. Widjaja, P. Bosch and H. La Roche, "Comparison of MME Signaling Loads for Long-Term-Evolution Architectures," *Vehicular Technology Conference Fall (VTC 2009-Fall), 2009 IEEE 70th*, Anchorage, AK, 2009, pp. 1-5.
- [10] 3GPP, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access," 3rd Generation Partnership Project (3GPP), TS 23.401, Jun. 2011.
- [11] D. Lam, D.C. Cox and J. Widom, "Teletraffic Modeling for Personal Communications Services," in *Communications Magazine, IEEE*, vol.35, no.2, pp.79-87, Feb 1997.